



PCT09

RAW SEQUENCE LISTING

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:52:10

Input Set : A:\NIELSEN4.txt

Output Set: N:\CRF4\01312003\I743414A.raw

3 <110> APPLICANT: Nielsen, Jens
 4 Nissen, Torben L
 5 Kielland-Brandt, Morten C
 7 <100> TITLE OF INVENTION: Metabolically engineered microbial cell with an altered
 8 metabolite production
 10 <140> FILE REFERENCE: NIELSEN=4
 12 <140> CURRENT APPLICATION NUMBER: 09/743,414A
 13 <141> CURRENT FILING DATE: 2001-01-10
 15 <150> PRIOR APPLICATION NUMBER: PCT/DK99/00397
 16 <151> PRIOR FILING DATE: 1999-07-12
 18 <170> PRIOR APPLICATION NUMBER: PA 149: 00967
 19 <151> PRIOR FILING DATE: 1999-07-10
 21 <160> NUMBER OF SEQ ID NOS: 14
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <210> LENGTH: 1395
 27 <210> TYPE: DNA
 28 <210> ORGANISM: Azotobacter vinelandii
 29 <210> FEATURE:
 31 <211> NAME/KEY: CDS
 32 <212> LOCATION: (1)..(1395)
 33 <213> OTHER INFORMATION:

W--> 35 <400> 1

36 atg gct gta tat aas tac gat gtg gtg gta atc ggc aca ggc cct gct 48
 W--> 37 Met Ala Val Tyr Xaa Tyr Asp Val Val Val Ile Gly Thr Gly Pro Ala
 38 1 5 10 15
 40 ggc aac ggc gaa ggc atg aat gcc gtg aag gcc ggc cgc aag gta ggc 96
 41 Gly Gln Gly Ala Ala Met Asn Ala Val Lys Ala Gly Arg Lys Val Ala
 42 20 25 30
 44 gtc ctg gat gat cgc ccc cag gtc gcc gcc aac tgc acc cac ctg gga 144
 45 Val Val Asp Asp Arg Pro Gln Val Gly Gly Asn Cys Thr His Leu Gly
 46 35 40 45
 48 aac att ccc tcc aag ggc ctg cgc cac tgg gtg cgg cag atc atg cag 192
 49 Thr Ile Pro Ser Lys Ala Leu Arg His Ser Val Arg Gln Ile Met Gln
 50 50 55 60
 51 tat aac aac aat cag ctg ttc cgc cag atc gcc gag cgc cgc tgg ttt 240
 52 Tyr Asn Asn Asn Pro Leu Phe Arg Gln Ile Gly Gln Pro Arg Trp Phe
 53 65 70 75 80
 55 tcc ttc gcc gat gtc ctg aag agc gcc gag cag gtc atc gcc aag cag 288
 56 Ser Phe Ala Asp Val Leu Lys Ser Ala Gln Gln Val Ile Ala Lys Gln
 57 85 90 95
 58 gtc tcc tcc cgc acc ccc tac tat ggc cgc aac ggc atc gat acc ttc 336
 59 Val Ser Ser Arg Thr Gly Tyr Tyr Ala Arg Asn Arg Ile Asp Thr Phe

RAW SEQUENCE LISTING

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:52:15

Input Set : A:\NIELSEN4.txt

Output Set : N:\CRF4\01312003\I743414A.raw

```

63      100      105      110
64 ttc ggg acc ggg agc ttc ttc gac gar cac acc atc gag gtc gtc cac 384
65 Phe Gly Thr Ala Ser Phe Cys Asp Glu His Thr Ile Glu Val Val His
66      115      120      125
68 ctg aac ggt atg gtc gaa aag ctg gtg gcc aag cag ttc gtc atc gag 432
69 Leu Asn Gly Met Val Glu Thr Leu Val Ala Lys Gln Phe Val Ile Ala
70      130      135      140
71 acc gga tgg cgt cca tac ccc ccg gcc gat gtc gat ttc acc cat cag 480
72 Thr Gly Ser Arg Pro Tyr Arg Pro Ala Asp Val Asp Phe Thr His Pro
73 145      150      155      160
75 cgg atc tac gac agc gac acc atc ctg agc ctg ggc cac acc ccg ctc 528
76 Arg Ile Tyr Asp Ser Asp Thr Ile Leu Ser Leu Gly His Thr Pro Arg
77      165      170      175
80 cgg ttc atc atc tac gga gag ggg gtg atc ggc tgc gaa tat gcc ttc 576
81 Arg Leu Ile Ile Tyr Gly Ala Gly Val Ile Gly Cys Glu Tyr Ala Ser
82      180      185      190
84 atc ttc agt ggg ctg ggt gtg ctg gtc gac ctg atc gac aac cgg gac 624
85 Ile Phe Ser Gly Leu Gly Val Leu Val Asp Leu Ile Asp Asn Arg Arg
86      195      200      205
88 cac ctg ctg agt ttc ctg gac gac gaa atc tcc gac tgg ctg agc ttc 672
89 Gln Leu Leu Ser Phe Leu Asp Asp Glu Ile Ser Asp Ser Leu Ser Tyr
90      210      215      220
92 cac ctg cgg aac aac aac gtc ctg atc cgg cac aac gar gaa tac gag 720
93 His Leu Arg Asn Asn Asn Val Leu Ile Arg His Asn Glu Glu Tyr Glu
94 225      230      235      240
96 cgt gtc gaa ggc ctg gac aac ggg gtg atc ctg cac ctg aag tcc ggc 768
97 Arg Val Glu Gly Leu Asp Asn Gly Val Ile Leu His Leu Lys Ser Gly
98      245      250      255
100 aag aag atc aag ggc gac gcc ttc ctg tgg agc aac ggc cgt acc agc 816
101 Lys Lys Ile Lys Ala Asp Ala Phe Leu Trp Ser Asn Gly Arg Thr Gly
102      260      265      270
104 aat aac gac aag ctg gcc ctg gag aac atc ggt ctg aag gcc aat agt 864
105 Asn Thr Asp Lys Leu Gly Leu Glu Asn Ile Gly Leu Lys Ala Asn Gly
106      275      280      285
108 cgc gaa cag atc cag gtc gac gag cac tac cgt acc gaa gtc agc aac 912
W--> 109 Arg Gly Gln Ile Gln Val Asp Glu His Tyr Arg Thr Glu Val Ser Xaa
110      290      295      300
112 att tat gcc gct ggt gac gtg atc ggc tgg ccg agc ctg gcc agc gcc 960
113 Ile Tyr Ala Ala Gly Asp Val Ile Gly Trp Pro Ser Leu Ala Ser Ala
114 305      310      315      320
116 gcc tat gac cag ggt cgt tgg gcc gcc ggc agt atc acc gag aac gat 1008
117 Ala Tyr Asp Gln Gly Arg Ser Ala Ala Gly Ser Ile Thr Glu Asn Asp
118      325      330      335
120 agc tgg cgt ttc gtc gac gac ctg ccg acc gcc atc tac acc att ccg 1056
121 Ser Trp Arg Phe Val Asp Asp Val Pro Thr Gly Ile Tyr Thr Ile Pro
122      340      345      350
124 gag atc agt tgg gtc gcc aag acc gag cgc gaa ctg acc cag gcc aag 1104
125 Glu Ile Ser Ser Val Gly Lys Thr Glu Arg Glu Leu Thr Gln Ala Lys
126      355      360      365

```

RAW SEQUENCE LISTING

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:12:10

Input Set : A:\NIELSEN4.txt

Output Set: N:\CRF4\01312003\I743414A.raw

```

128 gtt ccc tac gag gtc ggc aag gcc ttc ttc aag ggc atg gcc cgg gca 1152
129 Val Pro Tyr Glu Val Gly Lys Ala Phe Phe Lys Gly Met Ala Arg Ala
130      378      375      390
131 aag atc gcc gtc gag aag gcc ggc atg atg aag atc ctc ttt cac cgc 1200
132 Gln Ile Ala Val Glu Lys Ala Gly Met Leu Lys Ile Leu Phe His Arg
133 385      390      395      400
134 gag acc cta gaa atc ctc gcc gtc cac tgc ttc gcc tat cag gcc tgc 1248
135 Glu Thr Leu Glu Ile Leu Gly Val His Cys Phe Gly Tyr Gln Ala Ser
136      405      410      415
137 gaa atc gtc cat atc gcc gag gcc atc atg aac cag aag gcc gag gcc 1296
138 Glu Ile Val His Ile Gly Gln Ala Ile Met Asn Gln Lys Gly Glu Ala
139      420      425      430
140 cat acc ctc aag tat ttc atc aac acc acc ttc aac tac cgg acc atg 1344
141 Asn Thr Leu Lys Tyr Phe Ile Asn Thr Thr Phe Asn Tyr Pro Thr Met
142      435      440      445
143 gcc gag gcc tac gag gtc gcc gcc tac gag ggt ctc aat cgg ctt ttt 1392
144 Ala Glu Ala Tyr Arg Val Ala Ala Tyr Asp Gly Leu Asn Arg Leu Phe
145      450      455      460
146 gaa 1395

```

```

147 110 > SEQ ID NO: 1
148 111 > LENGTH: 464
149 112 > TYPE: PRT
150 113 > ORGANISM: Anotobacter vinelandii
151 114 > FEATURE:
152 115 > NAME/KEY: misc_feature
153 116 > LOCATION: (5)..(5)
154 117 > OTHER INFORMATION: The 'Xaa' at location 5 stands for Lys, or Asn.
155 118 > FEATURE:
156 119 > NAME/KEY: misc_feature
157 120 > LOCATION: (304)..(304)
158 121 > OTHER INFORMATION: The 'Xaa' at location 304 stands for Lys, or Asn.
159 122 > SEQUENCE: 2

```

```

W--> 172 Met Ala Val Tyr Xaa Tyr Asp Val Val Val Ile Gly Thr Gly Pro Ala
173      1      5      11      15
174 Gly Gln Gly Ala Ala Met Asn Ala Val Lys Ala Gly Arg Lys Val Ala
175      20      25      30
176 Val Val Asp Asp Arg Pro Gln Val Gly Gly Asn Cys Thr His Leu Gly
177      35      40      45
178 Thr Ile Pro Ser Lys Ala Leu Arg His Ser Val Arg Gln Ile Met Gln
179      50      55      60
180 Tyr Asn Asn Asn Pro Leu Phe Arg Gln Ile Gly Glu Pro Arg Trp Phe
181      65      70      75      80
182 Ser Phe Ala Asp Val Leu Lys Ser Ala Glu Gln Val Ile Ala Lys Gln
183      85      90      95
184 Val Ser Ser Arg Thr Gly Tyr Tyr Ala Arg Asn Arg Ile Asp Thr Phe
185      100      105      110
186 Phe Gly Thr Ala Ser Phe Cys Asp Glu His Thr Ile Glu Val Val His
187      115      120      125
188 Leu Asn Gly Met Val Glu Thr Leu Val Ala Lys Gln Phe Val Ile Ala

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/743,414A

DATE: 01/31/2003

TIME: 11:12:11

Input Set : A:\NIELSEN4.txt

Output Set : N:\CRF4\01312003\I743414A.raw

```

197      199      201      240
199 Thr Gly Ser Arg Pro Tyr Arg Pro Ala Asp Val Asp Phe Thr His Pro
200 145      150      155      160
201 Arg Ile Tyr Asp Ser Asp Thr Ile Leu Ser Leu Gly His Thr Pro Arg
202      165      170      175
203 Arg Leu Ile Ile Tyr Gly Ala Gly Val Ile Gly Cys Glu Tyr Ala Ser
204      180      185      190
205 Ile Phe Ser Gly Leu Gly Val Leu Val Asp Leu Ile Asp Asn Arg Asp
206      195      200      205
207 Gln Leu Leu Ser Phe Leu Asp Asp Gln Ile Ser Asp Ser Leu Ser Tyr
208      210      215      220
209 His Leu Arg Asn Asn Asn Val Leu Ile Arg His Asn Glu Glu Tyr Glu
210 225      230      235      240
211 Arg Val Glu Gly Leu Asp Asn Gly Val Ile Leu His Leu Lys Ser Gly
212      245      250      255
213 Lys Lys Ile Lys Ala Asp Ala Phe Leu Trp Ser Asn Gly Arg Thr Gly
214      260      265      270
215 Asn Thr Asp Lys Leu Gly Leu Glu Asn Ile Gly Leu Lys Ala Asn Gly
216      275      280      285
W--> 226 Arg Gly Gln Ile Gln Val Asp Glu His Tyr Arg Thr Glu Val Ser Xaa
217      290      295      300
218 Ile Tyr Ala Ala Gly Asp Val Ile Gly Trp Pro Ser Leu Ala Ser Ala
219 305      310      315      320
220 Ala Tyr Asp Gln Gly Arg Ser Ala Ala Gly Ser Ile Thr Glu Asn Asp
221      325      330      335
222 Ser Trp Arg Phe Val Asp Asp Val Pro Thr Gly Ile Tyr Thr Ile Pro
223      340      345      350
224 Glu Ile Ser Ser Val Gly Lys Thr Glu Arg Glu Leu Thr Gln Ala Lys
225      355      360      365
226 Val Pro Tyr Gln Val Gly Lys Ala Phe Phe Lys Gly Met Ala Arg Ala
227      370      375      380
228 Gln Ile Ala Val Glu Lys Ala Gly Met Leu Lys Ile Leu Phe His Arg
229 385      390      395      400
230 Glu Thr Leu Gln Ile Leu Gly Val His Cys Phe Gly Tyr Gln Ala Ser
231      405      410      415
232 Glu Ile Val His Ile Gly Gln Ala Ile Met Asn Gln Lys Gly Glu Ala
233      420      425      430
234 Asn Thr Leu Lys Tyr Phe Ile Asn Thr Thr Phe Asn Tyr Pro Thr Met
235      435      440      445
236 Ala Glu Ala Tyr Arg Val Ala Ala Tyr Asp Gly Leu Asn Arg Leu Phe
237      450      455      460
261 <10> SEQ ID NO: 1
262 <11> LENGTH: 52
263 <12> TYPE: DNA
264 <13> ORGANISM: Artificial Sequence
265 <20> FEATURE:
267 <23> OTHER INFORMATION: Description of Artificial Sequence: Primer
268 <400> SEQUENCE: 3
270 attcatgat gaattctatc ttatgggtccc attctttact gaactgatta ca 52

```

RAW SEQUENCE LISTING

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:52:10

Input Set : A:\NIELSEN4.txt

Output Set: N:\CRF4\01312003\I743414A.raw

```

273 <210> SEQ ID NO: 4
274 <211> LENGTH: 11
275 <212> TYPE: DNA
276 <213> ORGANISM: Artificial Sequence
278 <220> FEATURE:
279 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
281 <400> SEQUENCE: 4
282 ggcctact catatcaag catctctctcg ctggtaatt ttctgtctc ttgtctatca 60
283 gacttggaa g                                     71
286 <210> SEQ ID NO: 5
287 <211> LENGTH: 14
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
294 <400> SEQUENCE: 5
295 ggaattc agtctcaaa agag                                     24
298 <210> SEQ ID NO: 6
299 <211> LENGTH: 22
300 <212> TYPE: DNA
301 <213> ORGANISM: Artificial Sequence
303 <220> FEATURE:
304 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
306 <400> SEQUENCE: 6
307 ggaagatc ttctgaatg ctttttgata acaaaaaat                                     38
310 <210> SEQ ID NO: 7
311 <211> LENGTH: 18
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
318 <400> SEQUENCE: 7
319 ggaagatc tctgggaga gctaaaacga ttaacaaa                                     18
322 <210> SEQ ID NO: 8
323 <211> LENGTH: 19
324 <212> TYPE: DNA
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
330 <400> SEQUENCE: 8
331 ggcacacaac aaggtccta                                     19
334 <210> SEQ ID NO: 9
335 <211> LENGTH: 38
336 <212> TYPE: DNA
337 <213> ORGANISM: Artificial Sequence
339 <220> FEATURE:
340 <221> OTHER INFORMATION: Description of Artificial Sequence: Primer
342 <400> SEQUENCE: 9
343 gggggggatc ctctagaatg ccagtgttga aatcagac                                     38

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/743,414A

DATE: 1/31/2003
TIME: 11:52:11

Input Set : A:\NIELSEN4.txt
Output Set : N:\CRF4\01312003\I743414A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 5,304
Seq#:2; Xaa Pos. 5,304

VERIFICATION SUMMARY

DATE: 01/31/2003

PATENT APPLICATION: US/09/743,414A

TIME: 11:52:11

Input Set : A:\NIELSEN4.txt

Output Set: N:\CRF4\01312003\I743414A.raw

L:35 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1, line#:33
L:37 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:48
L:109 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:912
L:172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
L:226 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:288